ChinaXiv合作期刊 图书情报工作
第65卷 第24期 2021年12月

美国高校科学数据使用协议应用情况调查与启示*

■ 陈越[1] 都平平[2] 周琼[1]
[1] 中国矿业大学信息化建设与管理处 徐州 221116 [2] 中国矿业大学图书馆 徐州 221116

摘 要：[目的/意义]数据使用协议是连接科学数据供给方和使用方的纽带，是保障双方权益的基础。了解其概念、内容及应用情况，可为国内高校开展相关工作提供参考。[方法/过程]选取30所美国高校，运用网络调查法和文本分析法对数据使用协议相关政策、措施进行分析，总结其应用特色，然后对我国高校开展相关工作提出建议。[结果/结论]所调查美国高校普遍设置专门机构负责数据使用协议事宜，围绕协议的起草、审批和落地应用建立了较为完善的支撑保障体系。我国高校应规范数据使用协议的主体内容，加强协议与法律法规的衔接，采用多种途径构建高效的服务保障体系，形成切合实际的数据使用协议应用环境。

关键词：科学数据 数据使用协议 数据管理
分类号：G203
DOI：10.13266/j.issn.0252-3116.2021.24.014

科学数据共享可以促进数据重用和数据价值的充分发挥。随着科研经费投入增加与创新能力提升，我国已成为科学数据产出大国[1]。例如，由中国科学院、教育部、工信部等下属37个研究机构共同建设的国家基础学科公共科学数据中心，其科学数据资源覆盖物理、化学、材料等八大学科，有效促进数据综合利用和多学科交叉融合创新[2-3]。截至2021年4月，国家基础学科公共科学中心已集成数据资源1.515PB[2-3]。该中心管理的数据也被其他国家广泛下载和利用，来自美国、乌克兰等7个国家的累计数据访问与下载量已经超过25TB，仅来自美国的数据访问下载量就达到16TB[5]。

国家高度重视科学数据，将其作为基础性战略资源[3]。2018年3月国务院办公厅正式印发《科学数据管理办法》，为我国科学数据管理工作确定了行动纲领，为推动数据共享提供了政策依据[3]。但在具体操作层面，数据使用过程中的政策体系尚不完善，导致科研人员在数据共享与复用的具体实践中面临着一定困难：一方面，作为数据提供方，由于缺少规范数据使用的政策与法律框架，他们顾虑在提供数据时自身利益无法得到保护，并担心可能卷入因使用方对数据的误用、曲解而引发伦理和诉讼问题[6-8]；另一方面，作为共享数据的意向使用方，在申请数据遇到障碍时缺少相应支持渠道[9]。此外，从研究机构视角看，由于缺少对数据使用的管控，一些敏感数据的任意外流会引发数据安全隐患[10]。

数据供给与使用是相辅相成的。为推动数据开放共享，需要通过一定规约明确数据供给方与使用方之间的利益关系。数据使用协议（Data Use Agreement，DUA）作为连接供给方与使用方的纽带，能够对双方权责进行调整和约束，是实现数据共享利用的基础保障。与数据开放许可相比，数据使用协议具有特殊性：①数据使用协议覆盖的数据主要是涉及敏感信息的受限数据；②数据使用协议在数据提供方与使用方之间一对一签署，有时需要就条款进行协商；③数据使用协议在考虑提供方与使用方的权益之外，还需考虑与数据相关的其他个体权益。

在谷歌学术中以“data use agreement”作为关键字，设置搜索起止时间为2016-2020年，当限定搜索域名范围为美国教育机构（在检索式中加入“site：edu”）时，返回333条记录；而将搜索域名范围限定为欧洲（在检索式中加入“site：eu”）时，仅返回11条记录。对搜索结果进行综合分析，选择以部分美国高校为对象，调查其数据使用协议的应用情况，包括协议文本、相关

* 本文系国家社会科学基金项目“机构知识资源领域研究数据的关联组织和开放复用研究”（项目编号：19BTQ073）和江苏省现代教育技术智慧校园专项课题“智能时代高校人工智能通识教育培养模式研究”（项目编号：2020-R-84368）研究成果之一。
作者简介：陈越，副处长，工程师，硕士；都平平，副馆长，研究馆员，博士，通讯作者，E-mail：ppdu@cumt.edu.cn；周琼，工程师，硕士。
收稿日期：2021-03-29 修回日期：2021-08-28 本文起止页码：135-142 本文责任编辑：易飞

机制、支撑措施等,总结经验与特色,为我国高校完善科学数据使用相关的政策机制提供借鉴。

1 研究现状

当前,政府和科研机构广泛开展开放数据行动,推动了数据共享领域政策研究。在数据开放许可方面,国内研究者对许可类型、应用情况进行了较为详尽的调研分析[11-14],设计了适合我国特点的许可协议框架[15]。在数据使用协议方面,目前研究较少且分散,多数研究针对政府数据开放平台所使用协议的合规性进行分析[16],仅有部分研究在整体政策研讨中涉及到了使用协议,如将数据使用协议纳入高校数据安全保护框架[17],利用数据使用协议保障数据引用[18]等。

国外相关研究大致集中在 3 个方面:①较早的研究侧重于分析数据使用协议的合规要求,如从软件需求工程角度讨论数据使用协议的合规内容[19],通过案例分析发现过多的合规限制可能导致遵守协议与其他要求冲突[20]等;②随着数据使用协议逐渐被接受,相关研究转移到协议在不同场景下的应用,如对生物医学数据的存储和使用资格进行评估[21],在协议中通过设置双重认证来防止基因数据泄露[22],利用协议解决不完全匿名化数据的使用问题[23],通过数据使用协议推动质性研究数据重复利用等[24];③近期研究则较多关注协议应用过程中的问题及解决方案,包括调查和分析协议执行中的障碍因素[25-26],通过定期优化消除协议对数据使用的不利影响[27],通过构建和使用包含固定基本条款的协议模板加快审批速度[28]等。

综上所述,现有研究以"点"为主,针对的是数据使用协议的特定问题,或在探讨特定场景数据共享问题时涉及到与数据使用协议有关的主题,尚未从"面"的角度,对数据使用协议的应用情况进行调查分析。由于美国高校较早开始应用数据使用协议,在网站上普遍提供了数据使用协议相关资料,故本文拟选取部分美国高校,对其科学数据使用协议的应用情况进行调查,尝试梳理协议的概念、核心要素与支撑环境,结合我国实际对高校科学数据使用协议的应用提出建议。

2 定义说明、调查方法和基本数据

本文中科学数据的定义与《科学数据管理办法》保持一致,即通过研究、试验等过程产生或观测、调查、检查等方式取得并用于科研活动的数据及其衍生数据[29]。由于国外常用"Research Data"一词表述,文中在引用国外文献资料时直译为"研究数据",含义与科学数据等同。

以 U. S. News 发布的 2021 年美国大学排名为基础,选取前 30 位大学进行调查。以"data use agreement"为关键字,以该大学域名后缀(如 mit. edu)为搜索域,使用谷歌搜索引擎查找该大学网站上相关信息,调查时间截止到 2021 年 3 月。由于加州理工学院网站上的数据使用协议信息需要通过该校专用 VPN 才能访问,无法获取到相关数据,故将其去除,按排序补入塔夫茨大学作为调查对象。

围绕数据使用协议及其应用中涉及到的学校审核机构、辅助措施等内容进行调查,对所获取的信息进行整理,详情见表 1。

3 数据使用协议的概念和应用范围

3.1 数据使用协议的基本概念

由于目前并无统一的权威定义,各高校对于数据使用协议的表述不尽相同。解析这些不同的表述,从法律地位、监管内容及主要作用 3 个方面对数据使用协议的基本概念总结如下:

(1)法律地位:数据使用协议是一种由数据供给主体和使用主体签署,具有法律约束力的合同文件[30-31];

(2)监管内容:数据使用协议用于规范数据使用主体的行为,包括数据的获取、转移、使用等[32-33];

(3)主要作用:数据使用协议要求研究者依规利用数据,从而避免触犯法律条款[34],或无法发布研究成果[35]。

根据调查数据,芝加哥大学和密歇根大学较为特别,二者在涉及数据使用协议的网页中同时提到了数据共享协议(Data Sharing Agreement)。不过,芝加哥大学含有"数据共享协议"一词的网页,其链接名称是"数据使用协议";密歇根大学虽然在网页中提到数据使用协议包括数据共享协议[36],但在由研究管理部门发布的数据使用协议培训文档中没有提及[37],且网站提供的协议模板标题是"数据使用协议"[38]。此外利用谷歌学术搜索"Data Sharing Agreement",返回结果远少于搜索"Data Use Agreement";同时在返回结果中,"Data Sharing Agreement"的使用语境基本等同于"Data Use Agreement"。基于上述分析,可以认为主流高校倾向于使用"Data Use Agreement"一词。

3.2 数据使用协议的应用范围

并非所有的数据使用行为都需要签署数据使用协议。由公共机构产生、收集或不受商业条款和专利权

陈越，郝平平，周琼. 美国高校科学数据使用协议应用情况调查与启示[J]. 图书情报工作,2021,65(24):135－142.

表 1　美国高校(U. S. News TOP30)数据使用协议应用概况

序号	学校	DUA 样本	DUA 帮助文档	签署 DUA 申请方式	审核机构	过程说明
1	普林斯顿大学	无	无	邮件	Office of Research Integrity and Assurance	无
2	哈佛大学	无	有	线上	Data Use Agreements Negotiating Offices	有
3	哥伦比亚大学	无	无	邮件	The Office of Research Compliance and Training	无
4	麻省理工学院	有	有	邮件	Office of Strategic Alliances & Technology Transfer	无
5	耶鲁大学	有	有	线上	Office of Research Administration	无
6	斯坦福大学	无	有	邮件	Industrial Contract Office Office of Research Administration Research Management Group	无
7	芝加哥大学	有	有	线上	University Research Administration Official	无
8	宾夕法尼亚大学	有	有	线上	Office of Research Services	无
9	约翰斯霍普金斯大学	无	有	线上	Johns Hopkins University Research Administration (JHURA)	无
10	西北大学	无	有	线上	Office for Research	无
11	杜克大学	无	有	邮件	Research Facilitation & Data Security	无
12	达特茅斯学院	有	有	邮件	Data & Analytic Core	有
13	布朗大学	无	有	线上	Office of Research Integrity	无
14	范德堡大学	无	无	线上	Sponsored Programs and Administration (SPA) office	有
15	莱斯大学	无	无	邮件	The Office of General Counsel and the Compliance Office	无
16	圣路易斯华盛顿大学	有	有	邮件	Joint Research Office for Contracts	无
17	康奈尔大学	无	无	邮件	Office of Sponsored Programs	无
18	圣母大学	有	有	邮件	Notre Dame Research	无
19	加州大学洛杉矶分校	无	有	邮件	TDG/CTS&SR	无
20	埃默里大学	有	无	邮件	Office of Technology Transfer	无
21	加州大学伯克利分校	无	有	邮件	Intellectual Property & Industry Research Alliances(IPIRA)	有
22	乔治敦大学	无	无	邮件	Office of Technology Commercialization	无
23	密歇根大学	有	有	线上	Office of Research and Sponsored Projects	无
24	南加州大学	有	无	邮件	Office for the Protection of Research Subjects (OPRS)	无
25	卡耐基梅隆大学	无	无	邮件	Office of Sponsored Programs	有
26	弗吉尼亚大学	无	无	线上	Office of Sponsored Programs	无
27	北卡罗来纳大学教堂山分校	无	有	线上	Office of the Vice Chancellor for Research	无
28	维克森林大学	无	无	邮件	The Office of the Associate Provost for Research and Scholarly Inquiry	无
29	纽约大学	无	有	邮件	Office of Sponsored Programs	有
30	塔夫茨大学	无	无	线上	Office of the Vice Provost for Research	无

等束缚可以自由利用的数据，使用时可遵循开放数据许可协议。但对于一定领域内的研究数据，由于涉及敏感、保密等问题，访问和利用时需要签署数据使用协议。高校在发布数据使用协议相关政策时，均规定了其涉及的数据集，但对具体数据集的定义则有所不同。一些高校选择以《健康保险携带与责任法案》(Health Insurance Portability and Accountability Act，以下简称“HIPAA”)[39]中定义的有限数据集(Limited Data)为特定数据集[30, 34, 40－41]，但亦有高校将这一范围扩大到其他研究主题数据[33, 42]，或将所有敏感、受限的研究数据纳入其中[32, 43]。综合上述对于数据集的认定，需要纳入数据使用协议管理的特定数据集主要包括：

(1)健康信息数据(需要去除个人信息标识，如姓名、车牌、社保号码等)[40, 44]；

(2)未包含明确授权文件的商业数据，如金融、经济和商业信息[42]；

(3)其他被政府或资助机构认定为敏感或需要保密的研究数据，如用于心理学研究的问卷或访谈数据[24]。

4　数据使用协议的主要内容与参与主体

4.1　数据使用协议的主要内容

在所调研的高校中，有 10 所公开了数据使用协议样本。本研究采用 Nvivo 软件对所收集的数据使用协

议样本文件进行编码。一级编码采用开放式编码，逐条分析并抽取要点；接着对一级编码进行分析，发现其中的关联并归纳，最终将协议要点归为总体规定、数据相关权利、数据使用要求、协议有效期与终止、使用情况核查与报告、数据相关责任、其他约定等 7 个二级编码。编码层次如图 1 所示，图中色块的面积与参考点和编码项数量正相关，参考点和编码项数量标注方式为“参考点数量（编码项数量）”。

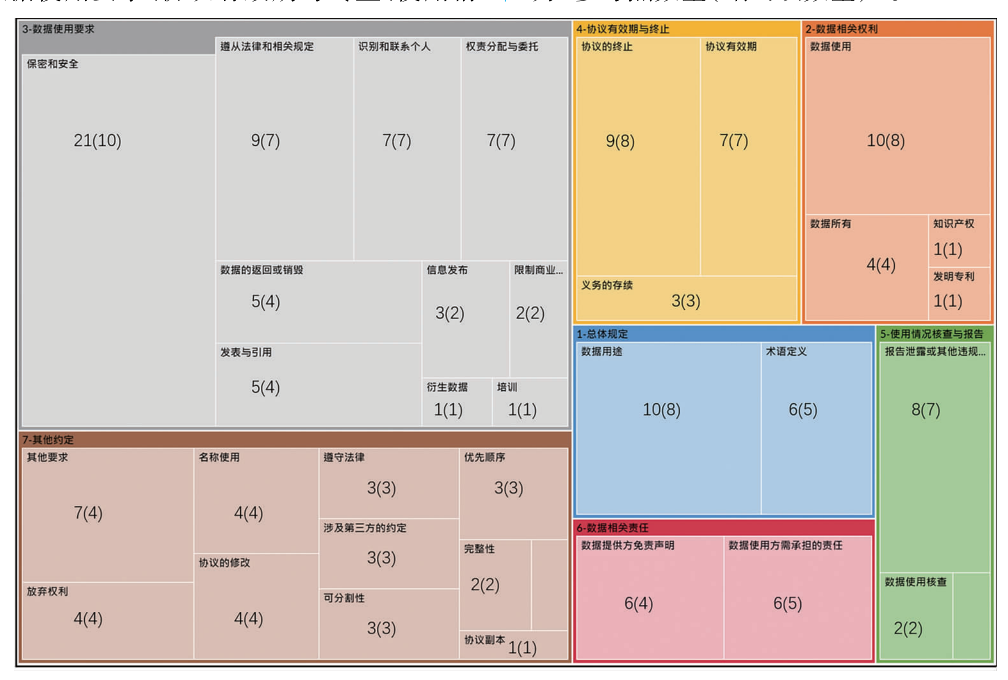


图 1　数据使用协议要点编码层次

4.1.1　总体规定

总体规定中包含术语定义和数据用途。通过对协议中的术语进行定义，签署各方可以知晓术语在协议上下文中的特定含义。例如“数据”中包括试验数据[45]，“有限数据集”采用《个人可识别健康信息隐私标准》中的定义[46]，“HIPAA”为《健康保险携带和责任法案》[47]。数据用途则在协议中限定数据只能用于特定的项目，项目名称须在协议正文或附件中列明[48]。

4.1.2　数据相关权利

协议一般规定，数据提供方拥有数据的所有权；同时，规定数据使用方拥有在限定数据用途下将数据用于研究的权利[46, 49]。

4.1.3　数据使用要求

本类所包含的要点是协议的核心，对数据使用方在使用过程中的行为提出了一系列要求。例如，在数据获取方面，要求数据必须用于特定项目[49]，新增项目即使使用同样数据也要重新申请[45]；在数据管理方面，数据使用方需要遵守相应法律，不能以泄露、分发或出售等形式将数据转移给第三方[45, 49]，但依法披露的情况除外[45, 48]。协议普遍强调数据安全与保密要求，并针对特定数据在使用过程中的安全要求加以细化，例如在隐私保护方面，强调不能利用获取到的数据来定位和联系特定人员，并以专门的条款来约定[45, 48, 50－51]。

4.1.4　协议有效期与终止

不同协议模板对有效期的约定差异较大，比较长的有 10 年[50]，较短的有 3 年或 1 年[52]，有效期可根据研究项目执行情况续延[45, 49]。也有协议规定有效期直到所有数据被退回提供方或销毁[46]，或根据条款终止为止[47]。显然，由于项目的多样性，协议签署过程中各方需要考虑有效期这一重要事项[53]。

为保护双方权益，多数协议模板都包含协议终止条款，用于约定协议终止的条件或场景，一般包括：①有效期满且未续签；②任何一方提出提前终止协议[45, 48－50, 52]。此外为及时阻止违反数据使用协议的行为，一些协议模板规定如果数据提供方认为使用方存在严重违约，可以提出终止协议，并在通知送达另一方后立即生效[45－47]。

4.1.5　使用情况核查与报告

为监管数据的使用，协议中可约定数据提供方可以对数据使用方进行检查，以监控遵守协议的情况[47]。

多数协议规定，如果在数据使用过程中发生泄露或其他违反协议约定的情况，使用方须在规定时限内

报告[45－47, 49, 51－52]，有时还要求提交包含补救措施的说明[46]。部分协议模板中约定数据使用方需要按照提供方要求提交数据使用报告，使用报告可以在研究完成时提交；即使研究未达到目的，也应报告数据提供方[45]。

4.1.6 数据相关责任

在这一类别的条款中，一般规定如果因使用方违约行为导致数据提供方遭受损失，则应承担相应的辩护、赔偿和追索等责任，使数据提供方免受损失[45－47, 51]。也有部分协议列入了数据提供方的免责声明，对于因使用数据造成的损失或索赔，除法律规定的情况外，提供方不承担责任[46]。

4.1.7 其他约定

除前述各项内容之外，协议模板中还包含了一些通用条款，如协议双方名称的使用约定[47－50]、协议适用法律、协议的修改及可分割性等[45－46, 48]。

4.2 数据使用协议的相关参与主体及相互关系

在数据使用协议的应用场景中，高校既可作为数据提供方，也可作为数据使用方。同时，数据提供方与使用方也可以是校外机构，如研究所、企业研发机构、医院等。

在数据使用方与提供方内部，存在两个与数据共享使用有关的主体：管理服务机构与研究人员。被调查高校都设立了专门机构负责数据使用协议相关事宜，提供协议的咨询与审核服务。数据使用协议的审核往往涉及校内多个部门，有时还需要与校外机构就具体条款进行商谈[31－32]。这些校外机构既有可能是数据的供给方，也可能是数据的潜在使用方。研究人员则按照相关工作要求，填写申请及起草协议，并在协议签署后，执行具体的数据提供或数据获取工作。围绕数据使用协议起草、签署与执行过程，数据提供方与使用方相关主体的业务关系如图2所示：

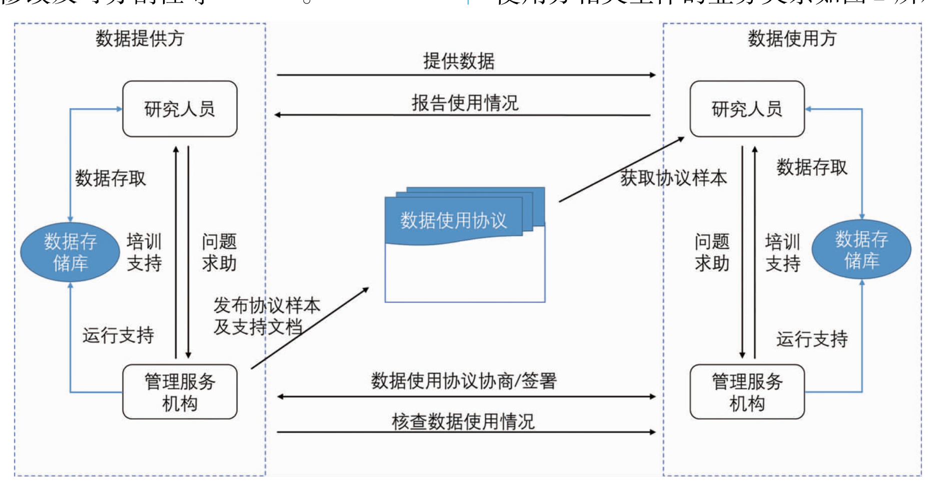


图2 数据使用协议涉及的各主体之间的关系

5 数据使用协议配套的支撑保障措施

由于数据使用协议涉及多方主体，管理机构不仅要进行文本审查，还需要代表高校就相关条款与校内外各方进行谈判，其审批过程较为复杂[42]，一些高校给出的时间从数周到数月不等[54]。一项访谈研究显示，83%的被访机构表示收到过教师对于执行数据使用协议的抱怨，主要意见焦点是审批延时[25]。许多高校也认识到，良好的沟通有助于教师了解预期时间，更好地安排研究计划，为此对数据使用协议的审批步骤给出了较完善的说明[30－31, 42, 55]，一些学校提供了流程图[56－58]。此外，围绕数据使用协议的起草、审批和落地应用，还有高校建立了一系列的支撑保障措施。

5.1 提供多种形式的说明及培训材料

由于监管合规的要求，研究人员在准备数据使用协议时，需要考虑更多，例如：哪些数据需要申请签署使用协议？要准备哪些材料？如何使用在线提交/审批系统？为给研究人员提供帮助，除在网站中设置问答栏目介绍基础概念和常见问题外，一些高校还采取了新的举措。斯坦福大学设计了数据使用协议决策树，协助研究人员判断何时需要协议以及向谁提交协议申请[59]；密歇根大学制作了快速参考卡片，将在线审批系统的常见问题、审批步骤、联系方式以简短清晰的形式提供给研究人员[60]；弗吉尼亚大学录制了在线提交系统的使用视频，指导用户一步步完成数据使用协议的申请提交过程[61]。

5.2 提供网络化的数据使用协议审批环境

调研的30所高校中已有12所高校采用线上方式提交数据使用协议申请表。还有部分高校建立了管理系统，将协议的申请、审查与批准过程全部转到线上。如约翰斯霍普金斯大学建立了协议工作流系统

(JAWS)[62];芝加哥大学、宾夕法尼亚大学在其研究管理系统中设立专门的功能模块[41,63],用于教师提交与研究相关的协议,在工作人员进行线上审核的同时,也便于教师利用系统跟踪批准过程。

5.3 多部门合作为数据使用协议提供支持服务

数据使用协议涉及知识产权、数据安全、隐私保护、研究伦理等多个方面,一些高校从整合科研数据管理与安全服务角度出发,建立跨部门工作机制协调处理数据使用协议相关事宜。如普林斯顿大学由图书馆、教务部门、信息技术部门组建研究数据服务联合工作组,将使用管理作为数据生命周期管理的一部分[64];范德堡大学信息技术部门与研究管理部门密切合作,针对数据使用协议涵盖的数据,参与协调实施信息安全防护措施[32];杜克大学通过跨部门研究计划办公室为研究生提供受保护的数据网络服务,研究数据安全团队负责根据数据使用协议或数据提供方的要求销毁数据[65]。

6 我国高校推进数据使用协议相关工作的思考

6.1 规范数据使用协议的主体内容,加强协议与相关法律法规的衔接

参考国外高校经验,应在学校层面统一数据使用协议模板,围绕数据使用要求、数据相关权利、有效期与终止条件等协议核心要素,规范协议的主体内容,避免因协议起草、审核人员领域知识欠缺产生规则短板与法律风险。

加快落实法律法规要求,确保协议合规签署和执行。一是根据《数据安全法》第二十一条[66-67]和《科学数据管理办法》第二十条[29]的要求,尽快开展本单位科学数据分级分类工作,建立科学数据目录,明确不同数据集的范围、保护级别、开放程度,以便在准备数据使用协议时准确引用。二是按照《个人信息保护法》的要求,对本单位已经签署的数据使用协议和拥有的科学数据集进行审核;对未达到法律要求的协议及时协商修订或终止,对含有法律规定的个人敏感信息的数据集进行去标识化或加密处理[68],确保数据使用协议依法执行。

6.2 建立高效的服务保障体系,推动数据使用协议落地应用

美国高校普遍重视数据使用协议服务支撑环境的建设,调研高校均设立了专门机构负责模板管理、政策咨询、培训支持、协商审核等工作,是协议的支撑主体。尽管如此,仍有调查显示,由于数据使用协议条款不断增长、合规要求日趋复杂以及跨部门协商场景的增加,从准备协议到交付数据的拖延依然很严重[25]。

有鉴于此,我国高校数据使用协议在落地应用的过程中,应以提高效能作为推进工作的首要考虑:①在学校层面就科学数据使用管理建立跨部门的协同机制,打通从政策宣传、知识培训到协议起草审核的业务链条,整合图书馆、科研管理部门、信息化部门等多方资源,建立以数据使用为主题的一站式服务入口,提供与数据使用协议有关的政策、模板、问答、培训文档等资源,并与各部门相关业务页面互链,便于研究人员快速定位到关心的问题,提高数据使用协议相关信息的获取效率;②依托信息技术全面提升服务能力。以数据使用协议审批过程的线上化为起点,逐步引入在线协同编辑、协议电子签章等功能。未来还可建立与数据使用协议相关的知识库,结合自然语言处理、机器学习等人工智能技术,尝试开展智能问答、协议文本智能审核[69]等业务,提高咨询与审批服务效率;③在推广数据使用协议时,要避免以文件说明文件的习惯做法,借鉴美国高校经验,采取多样化的宣传培训形式,充分利用公众号、短视频等新媒体平台,提升数据使用协议的宣传推广成效。

7 结语

党的十九届五中全会提出,数据要作为一种要素参与市场化配置,发挥更大效用。推动数据使用协议的落地应用,不仅仅是某个具体主管部门的问题,更考验的是高校的整体治理能力,需要在政策、机制、技术、服务等多方面做好准备,在实践中不断思考和探索,以应对不断发展变化的科学数据需求。

参考文献:

[1] 陈传夫, 李秋实. 数据开放获取使科学惠及更广——中国开放科学与科学数据开放获取的进展与前瞻[J]. 信息资源管理学报, 2020,10(1): 4-13.
[2] 国家科技基础条件平台中心. 国家基础学科公共科学数据中心[EB/OL]. [2021-08-25]. https://www.escience.org.cn/data-center/detail? orgId=2ce818534b401c7fa2f2b28d816e5c73.
[3] 国家科技基础条件平台中心. 国家科学数据资源发展报告2019[M]. 北京: 科学技术文献出版社, 2020.
[4] 国家基础学科公共科学数据中心. 数据服务成效统计[EB/OL]. [2021-08-25]. http://www.nbsdc.cn/general/effectOverview.
[5] 基础科学数据共享网数据服务监控与统计系统[EB/OL]. [2021-02-17]. http://msis.nsdc.cn/report.action? gid=1.
[6] 盛小平, 武彤. 国内外科学数据开放共享研究综述[J]. 图书情报工作, 2019,63(17): 6-14.
[7] 王春晓, 陈姝彤, 徐坤. 研究生科学数据共享态度与共享意愿关系研究[J]. 情报科学, 2020,38(12): 78-84.
[8] 李斯. 图书情报科研人员对高校图书馆科学数据共享的感知风险研究[J]. 图书馆学研究, 2019(9): 44-53.

[9] 国家科技基础条件平台中心. 国家科学数据资源发展报告 2017[M]. 北京：科学技术文献出版社，2018.
[10] 李洋，温亮明. 我国科学数据外流：表现、问题与对策[J]. 图书馆杂志，2019,38(12)：72 – 81.
[11] 黄如花，李楠. 开放数据的许可协议类型研究[J]. 图书馆，2016(8)：16 – 21.
[12] 黄如花，李楠. 国外政府数据开放许可协议采用情况的调查与分析[J]. 图书情报工作，2016,60(13)：5 – 12.
[13] 杨敏，夏翠娟，徐华博. 开放数据许可协议及其在图书馆领域的应用[J]. 图书馆论坛，2016,36(6)：91 – 98.
[14] 范丽莉，张蕾予. 政府数据开放许可协议研究：现状调查与机制构建[J]. 渭南师范学院学报，2020,35(9)：48 – 55.
[15] 范佳佳. 中国政府数据开放许可协议(CLOD)研究[J]. 中国行政管理，2019(1)：23 – 29.
[16] 完颜邓邓，陶成煦. 政府数据开放平台用户协议合规性评估[J]. 图书馆论坛，2021,41(7)：116 – 124.
[17] 刘桂锋，阮冰颖，包翔. 数据生命周期视角下高校科学数据安全内容框架构建[J]. 情报杂志，2021,40(2)：146 – 153.
[18] 黄国彬，刘馨然，姜颖. 影响科学数据引用的外部因素分析[J]. 数字图书馆论坛，2017(6)：2 – 8.
[19] SCHMIDT J Y，ANTÓN A I，WILLIAMS L，et al. The role of data use agreements in specifying legally compliant software requirements[C]// Fourth international workshop on requirements engineering and law. Piscataway：IEEE，c2011：1 – 4.
[20] KANOUS A，BROCK E. Contractual limitations on data sharing [EB/OL]. [2021 – 03 – 18]. https://deepblue.lib.umich.edu/handle/2027.42/123016.
[21] TOGA A W，DINOV I D. Sharing big biomedical data[J]. Journal of big data，2015,2(1)：1 – 12.
[22] SAVAGE S R. Characterizing the risks and harms of linking genomic information to individuals[J]. IEEE security & privacy，2017,15(5)：14 – 19.
[23] CANHAM S. Long-term Management of data and secondary use [M]//PIANTADOSI S，MEINERT C L. Principles and practice of clinical trials. Cham：Springer，2020：1 – 30.
[24] DUBOIS J M，STRAIT M，WALSH H. Is it time to share qualitative research data? [J]. Qualitative psychology，2018,5(3)：380 – 393.
[25] MELLO M M，TRIANTIS G，STANTON R，et al. Waiting for data：barriers to executing data use agreements[J]. Science，2020，367(6474)：150 – 152.
[26] WIEHE S E，ROSENMAN M B，CHARTASH D，et al. A solutions-based approach to building data-sharing partnerships [J]. Egems，2018,6(1)：20.
[27] LEARNED K，DURBIN A，CURRIE R，et al. Barriers to accessing public cancer genomic data[J]. Scientific data，2019,6(1)：1 – 7.
[28] KIM M，SHEN J，WENTWORTH L. Data infrastructure for partnership research：structures and processes used in the stanford-sfusd partnership[EB/OL]. [2021 – 08 – 26]. https://cepa.stanford.edu/sites/default/files/20190516_technical_paper_for_publication.pdf.
[29] 国务院办公厅. 关于印发科学数据管理办法的通知[EB/OL]. [2021 – 08 – 21]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
[30] Data use agreements[EB/OL]. [2021 – 08 – 21]. https://researchdatamanagement.harvard.edu/data-use-agreements.
[31] Navigating dua[EB/OL]. [2021 – 08 – 22]. https://securedata.uchicago.edu/navigating-dua/.
[32] Data use agreements[EB/OL]. [2021 – 08 – 26]. https://it.vanderbilt.edu/services/catalog/security/security_policy_and_compliance/Data_Use_Agreements.php.
[33] Data use agreements[EB/OL]. [2021 – 08 – 20]. https://osr.northwestern.edu/dua/.
[34] Determine what is needed before a contract can be negotiated[EB/OL]. [2021 – 08 – 26]. https://research.wustl.edu/determine-what-is-needed-before-a-contract-can-be-negotiated/.
[35] Data use agreements and restricted data[EB/OL]. https://researchdata.princeton.edu/research-lifecycle-guide/data-use-agreements-and-restricted-data.
[36] Unfunded agreement types[EB/OL]. [2021 – 08 – 25]. https://orsp.umich.edu/unfunded-agreement-types.
[37] WOODS P. Demystifying data use agreements[EB/OL]. [2021 – 08 – 25]. https://orsp.umich.edu/sites/default/files/resource-download/basics_of_data_use_agreements.pdf.
[38] Model data sharing agreement[EB/OL]. [2021 – 08 – 25]. https://deepblue.lib.umich.edu/handle/2027.42/136146.
[39] 刘抒悦，高上知，商瑾，等. 美国《健康保险携带和责任法案》中关于生物医学研究的规定及其影响[J]. 中国医学伦理学，2016,29(6)：1011 – 1014.
[40] Limited data set policy[EB/OL]. [2021 – 08 – 22]. https://universitypolicies.columbia.edu/content/limited-data-set-policy.
[41] Research inventory system[EB/OL]. [2021 – 08 – 22]. https://researchservices.upenn.edu/systems/research-inventory-system/.
[42] Obtaining data or software for research[EB/OL]. [2021 – 08 – 26]. https://ipira.berkeley.edu/data-and-software-agreements-research.
[43] Research facilitation & data security[EB/OL]. [2021 – 08 – 22]. https://researchinitiatives.duke.edu/programs-services/research-facilitation-data-security.
[44] Sample data use agreement[EB/OL]. [2021 – 08 – 21]. https://research.nd.edu/assets/204294/sample_data_use_agreement.docx.
[45] Yoda project quml dua march 2019[EB/OL]. [2021 – 08 – 26]. https://yoda.yale.edu/sites/default/files/files/YODA%20Project%20QMUL%20DUA%20March%202019.pdf.
[46] Template data use agreement outgoing[EB/OL]. [2021 – 08 – 27]. https://research.wustl.edu/wp-content/uploads/2017/07/Template-Data-Use-Agreement-Outgoing.pdf.
[47] Third party health data agreements[EB/OL]. [2021 – 08 – 27]. https://ctsi.ucla.edu/researcher-resources/files/view/docs/3rdparty/UCLA_Health_Third_Party_Health_Data_Agreements_-_Mandated_Contract_Terms_07102019.pdf.
[48] Data use agreement template[EB/OL]. [2021 – 08 – 27]. ht-

tps://research. nd. edu/assets/204108/data_use_agreement_template_non_hipaa_. docx.
[49] Outgoing dua for limited data set[EB/OL]. [2021-08-27]. https://uchicago. box. com/s/yi0rhvyech9nf0ljo4f4ggpqkp50z4i2.
[50] Mit-rfsleep_data_use_agreement[EB/OL]. [2021-08-26]. http://sleep. csail. mit. edu/files/rfsleep_data_use_agreement. pdf.
[51] PLISKE N. Data use agreement[EB/OL]. [2021-08-27]. https://irb. upenn. edu/sites/default/files/datauseagreement_0. doc.
[52] Fill in the blank dua[EB/OL]. [2021-08-27]. https://www. psc. isr. umich. edu/dis/data/restricted/rdc/Fill-in-the-Blank-DUA. pdf.
[53] Whatshould be addressed in a data-sharing agreement? [EB/OL]. [2021-08-27]. https://ura. uchicago. edu/page/data-sharing-agreements.
[54] Dua faqs[EB/OL]. [2021-08-23]. https://sites. dartmouth. edu/dac/dua/irb-and-dua/faqs/.
[55] Submittal to osp[EB/OL]. [2021-08-26]. https://www. cmu. edu/osp/contracts/contracts-process/submittal. html.
[56] Obtaining a dua[EB/OL]. [2021-08-26]. https://sites. dartmouth. edu/dac/dua/irb-and-dua/#Obtaining_a_DUA.
[57] Dua intake form[EB/OL]. [2021-08-28]. https://redcap. vanderbilt. edu/surveys/? s=KKEKR8PKRX.
[58] Non-funded agreements[EB/OL]. [2021-08-21]. http://www. nyu. edu/content/dam/nyu/research/documents/OSP/NFA-processmap. pdf.
[59] Dua decision tree[EB/OL]. [2021-08-20]. https://stanford. box. com/s/x38up0ydkw8i2c7udbeddw3ain295ejf.
[60] Pi & project team: proposal, subcontract, and unfunded agreement preparation[EB/OL]. [2021-08-20]. https://its. umich. edu/academics-research/research/eresearch/proposal-management/reference-materials/proposal-preparation.
[61] Using the enfa[EB/OL]. [2021-08-20]. https://info. researchuva. virginia. edu/using-enfa.
[62] Jhura agreement workflow system[EB/OL]. [2021-08-20]. https://research. jhu. edu/jhura/jaws/.
[63] Aura[EB/OL]. [2021-08-22]. https://aura. uchicago. edu/.
[64] Research data lifecycle[EB/OL]. [2021-08-24]. https://researchdata. princeton. edu/research-lifecycle-guide/research-lifecycle-guide.
[65] Protected research data network policies[EB/OL]. [2021-08-27]. https://ssri. duke. edu/sites/ssri. duke. edu/files/pdfs/PRDN%20Policies%202020. pdf.
[66] 中国人大网. 中华人民共和国数据安全法[EB/OL]. [2021-08-26]. http://www. npc. gov. cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a. shtml.
[67] 刘桂锋, 阮冰颖, 刘琼. 加强数据安全防护 提升数据治理能力——《中华人民共和国数据安全法(草案)》解读[J]. 农业图书情报学报, 2021,33(04): 4-13.
[68] 中国政府网. 中华人民共和国个人信息保护法[EB/OL]. [2021-08-26]. http://www. gov. cn/xinwen/2021-08/20/content_5632486. htm.
[69] 黄俏娟, 罗旭东. 人工智能与法律结合的现状及发展趋势[J]. 计算机科学, 2018,45(12): 1-11.

作者贡献说明：

陈越:论文撰写与修改；
都平平:研究框架设计与论文修改；
周琼:资料收集、内容分析与表格绘制。

Investigation and Enlightenment on the Application of Scientific Data Use Agreement in American Universities

Chen Yue[1] Du Pingping[2] Zhou Qiong[1]
[1] Division of Information Technology and Management, China University of Mining and Technology, Xuzhou 221116
[2] China University of Mining and Technology Library, Xuzhou 221116

Abstract: [**Purpose/significance**] As a link connecting scientific data providers and users, data use agreements (DUAs) play a fundamental role in protecting the rights and interests of both parties. A good understanding of its concept, content and application will provide references for work in this field in domestic universities and colleges. [**Method/process**] This research selected thirty American universities. By means of online survey and text analysis, this research analyzed policies and measures concerning DUAs and summarized features of their application, then put forward suggestions for Chinese universities to promote their work. [**Result/conclusion**] American universities under investigation have set up a special division for DUAs, and established a relatively complete support and guarantee system for the drafting, approval, and application of DUAs. Chinese universities should standardize the main contents of DUAs, strengthen the link-up between DUAs and laws and regulations, and develop a variety of ways to build an efficient service guarantee system, so as to create a practical application environment for DUAs.

Keywords: scientific data data use agreements data management